Testing nonparametric hypotheses for stationary processes by estimating minimal distances

Holger Dette, Tatjana Kinsvater, Mathias Vetter
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum
Germany
email: holger.dette@ruhr-uni-bochum.de
FAX: +49 2 34 32 14 559

May 4, 2010

Abstract

In this paper new tests for nonparametric hypotheses in stationary processes are proposed. Our approach is based on an estimate of the L^2-distance between the spectral density matrix and its best approximation under the null hypothesis. We explain the main idea in the problem of testing for a constant spectral density matrix and in the problem of comparing the spectral densities of several correlated stationary time series. The method is based on direct estimation of integrals of the spectral density matrix and does not require the specification of smoothing parameters. We show that the limit distribution of the proposed test statistic is normal and investigate the finite sample properties of the resulting tests by means of a small simulation study.

AMS subject classification: 62M15, 62G10
Keywords and phrases: spectral density, stationary process, goodness-of-fit tests, L^2-distance, integrated periodogram

1 Introduction

The problem of testing hypotheses about the second order properties of a multivariate stationary time series has found considerable attention in the literature. Many important hypotheses can be expressed in terms of functionals of the spectral density matrix. Several authors have proposed tests based on

the integrated periodogram [see Anderson (1993) or Chen and Romano (1999) among others]. Because on one hand, test statistics based on the integrated periodogram are usually not distribution free and, on the other hand, the type of hypotheses that can be tested by the integrated periodogram is limited, alternative methods have been proposed which are based on estimates of the spectral density [see Taniguchi and Kondo (1993), Taniguchi et al. (1996), Paparoditis (2000), Dette and Spreckelsen (2003), Eichler (2008) or Dette and Paparoditis (2009), among others]. These methods usually yield a normal distribution as the asymptotic law of the corresponding test statistics, but require the specification of a smoothing parameter in order to get consistent estimates of the spectral density matrix. As a consequence, the outcome of the testing procedure depends sensitively on this regularization. In the present paper we propose an alternative method for testing nonparametric or semiparametric hypotheses regarding the second order properties of stationary processes. Our approach is applicable to a broad class of hypotheses and is based on the estimation of the minimal L^2-distance between the spectral density matrix of a stationary time series and its best approximation in the class of all densities which satisfy the null hypothesis. As a consequence, it only requires estimates of the integrated spectral density matrix over the full frequency domain which are easily available by an appropriate summation of the periodogram. On one hand, this avoids the problem of smoothing the periodogram, and on the other hand, the limiting distributions [after an appropriate standardization] are asymptotically normally distributed, where the corresponding variance also contains only integrals of the components of the spectral density matrix over the full frequency domain and is thus easy to estimate.
In Section 2 we introduce the necessary notation, the basic assumptions and explain the main principle of our approach in the case of testing the null hypothesis of a white noise process. Section 3 is devoted to the problem of comparing spectral densities of a multivariate time series [see Eichler (2008) or Dette and Paparoditis (2009)]. In all cases we show that the proposed test statistic is asymptotically normally distributed, and a simple goodness-of-fit test for the null hypothesis is proposed, which uses the quantiles of the standard normal distribution. In Section 4 the finite sample properties of the new test procedures are investigated by means of a small simulation study, and some conclusions how the results can be extended to other testing problems are given in Section 5. Some technical details are deferred to an appendix in Section 6.

2 The main principle: testing for a white noise process

Let $\{\mathbf{X}_t\}_{t\in\mathbb{Z}}$ denote an m-dimensional stationary process with values in $\mathbb{R}^m$ which has a linear representation of the form

$$\mathbf{X}_t = (X_{1,t}, X_{2,t}, \dots, X_{m,t})^T = \sum_{j=-\infty}^{\infty} C_j \mathbf{Z}_{t-j} \qquad t \in \mathbb{Z}, \tag{2.1}$$

where $\{\mathbf{Z}_t\}_{t\in\mathbb{Z}} = \{(Z_{1,t}, Z_{2,t}, \dots, Z_{m,t})^T\}_{t\in\mathbb{Z}}$ denotes an m-dimensional Gaussian white noise process with covariance matrix

$$\Sigma = (\sigma_{r,s})_{r,s=1,\dots,m},$$

and such that the elements of the matrices $C_j = (c_j^{rs})_{r,s=1,\dots,m} \in \mathbb{R}^{m\times m}$ $(j \in \mathbb{Z})$ satisfy

$$\sum_{j\in\mathbb{Z}} |j||c_j^{rs}| < \infty, \qquad r,s = 1,2,\dots,m. \tag{2.2}$$

Throughout this paper we assume that the process $\{\mathbf{X}_t\}_{t\in\mathbb{Z}}$ has a spectral density matrix, say $f = (f_{ij})_{i,j=1}^m$, with elements f_{ij} which are Hölder continuous of order $L > 1/2$.

In order to explain the basic principle of our approach we consider the hypothesis that $\{\mathbf{X}_t\}_{t\in\mathbb{Z}}$ is a white noise process, that is

$$H_0 : f(\lambda) = \Sigma \qquad \text{versus} \qquad H_0 : f(\lambda) \neq \Sigma \tag{2.3}$$

for some (unknown) hermitian matrix $\Sigma \in \mathbb{R}^{m\times m}$. The problem of testing for white noise has a long history in statistics and econometrics with time series data. The most popular approach has been the Box-Pierce test [see Box and Pierce (1970)], which investigates the first p autocorrelations. Since this seminal work numerous authors have proposed alternative procedures for testing for white noise in stationary processes [see Ljung and Box (1978), Monti (1994) and Peña and Rodriguez (2002) among many others]. Many authors consider the problem of testing for white noise in an ARMA(p,q) process [see e.g. Mokkadem (1997) or Dette and Spreckelsen (2000)]. On the contrary, the problem of testing for white noise against general alternatives is more complicated. The classical approach involves the standardized cumulative periodogram [see e.g. Bartlett (1955) and Dahlhaus (1985)]. In this section we propose an alternative test for this problem, which is based on the concept of best L^2-approximation. For the construction of an appropriate test statistic we investigate the problem of approximating the true spectral density matrix $f(\lambda)$ by a constant function. A natural distance is given by

$$\begin{aligned} M^2(\Sigma) &= \int_{-\pi}^{\pi} \mathrm{tr}\,[(f(\lambda)-\Sigma)(f(\lambda)-\Sigma)^*]\,d\lambda \\ &= \int_{-\pi}^{\pi} \mathrm{tr}\,[(f(\lambda)-\Sigma_0)(f(\lambda)-\Sigma_0)^*]\,d\lambda + \int_{-\pi}^{\pi} \mathrm{tr}\,[(\Sigma-\Sigma_0)(\Sigma-\Sigma_0)^*]\,d\lambda, \end{aligned}$$

where the matrix $\Sigma_0 \in \mathbb{R}^{m\times m}$ is defined as

$$\Sigma_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\lambda)d\lambda,$$

and the identity above follows by a straightforward calculation. Consequently, we obtain for the minimum of the function $M^2(\cdot)$ on the set of all hermitian matrices

$$\begin{aligned} M^2 &= \min\{M^2(\Sigma) \mid \Sigma \in \mathbb{R}^{m\times m},\ \Sigma^* = \Sigma\} = M^2(\Sigma_0) \\ &= \mathrm{tr}\Big\{ \int_{-\pi}^{\pi} f(\lambda)f^*(\lambda)d\lambda - \frac{1}{2\pi}\Big(\int_{-\pi}^{\pi} f(\lambda)d\lambda \int_{-\pi}^{\pi} f^*(\lambda)d\lambda\Big)\Big\}. \end{aligned}$$

In order to estimate the minimal distance M^2, consider the periodogram

$$I_n(\lambda_j) = J_n(\lambda_j)J_n^*(\lambda_j), \quad J_n(\lambda_j) = \frac{1}{\sqrt{n}}\sum_{t=1}^n \boldsymbol{X}_t e^{-it\lambda_j}, \tag{2.4}$$

at the Fourier frequency $\lambda_j = \frac{2\pi j}{n} \in [-\pi, \pi]$ for any $j = -\lfloor \frac{n-1}{2} \rfloor, \dots, \lfloor \frac{n}{2} \rfloor$ and define the statistic

$$T_n = \frac{1}{2\pi} \mathrm{tr}\Big(T_{n,2} - T_{n,1} T_{n,1}^*\Big), \tag{2.5}$$

where the random variables $T_{n,1}$ and $T_{n,2}$ are given by

$$\begin{aligned} T_{n,1} &= \frac{1}{n} \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} \Big(I_n(\lambda_k) + \overline{I_n(\lambda_k)}\Big), \\ T_{n,2} &= \frac{2}{n} \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} I_n(\lambda_k) I_n^*(\lambda_{k-1}). \end{aligned}$$

Intuitively this makes sense, as we have with $E[I_n(\lambda_k)_{ij}] \approx 2\pi f_{ij}(\lambda_k)$ [note also that $\overline{f(\lambda)} = f(-\lambda) = f^T(\lambda)$ and corresponding relations for the periodograms hold]

$$\begin{aligned} E[T_{n,1}] &\approx \frac{2\pi}{n} \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} (f(\lambda_k) + \overline{f(\lambda_k)}) = \frac{2\pi}{n} \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} (f(\lambda_k) + f(-\lambda_k)) \approx \int_{-\pi}^{\pi} f(\lambda) d\lambda, \\ E[T_{n,2}] &\approx \frac{2}{n} \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} E[I_n(\lambda_k)] E[I_n^*(\lambda_{k-1})] \approx \frac{8\pi^2}{n} \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} f(\lambda_k) f^*(\lambda_k) \approx 4\pi \int_0^{\pi} f(\lambda) f^*(\lambda) d\lambda. \end{aligned}$$

Using the relation

$$2 \int_0^{\pi} \mathrm{tr}\ f(\lambda) f^*(\lambda) d\lambda = \int_{-\pi}^{\pi} \mathrm{tr}\ f(\lambda) f^*(\lambda) d\lambda$$

this calculation motivates (heuristically) the approximation $E[T_n] \approx M^2$ and indicates that T_n is a consistent estimate of the minimal distance M^2. Our first main result makes this heuristic argument rigorous and specifies the asymptotic distribution of T_n under the null hypothesis and the alternative.

Theorem 2.1 *If $\{\boldsymbol{X}_t\}_{t \in \mathbb{Z}}$ denotes a stationary process satisfying (2.1) and (2.2) with Hölder continuous spectral density matrix of order $L > 1/2$, then as $n \to \infty$*

$$\sqrt{n}(T_n - M^2) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \tau_{M^2}^2),$$

where the asymptotic variance is given by

$$\begin{aligned} \tau_{M^2}^2 &= 4\pi \int_{-\pi}^{\pi} \Big\{ 4\ \mathrm{tr}\ (f^4(\lambda)) + \Big(\mathrm{tr}\ (f^2(\lambda))\Big)^2 \Big\} d\lambda - 16 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \mathrm{tr}\ \Big(f^3(\lambda) f(\mu)\Big) d\lambda d\mu \\ &+ \frac{4}{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \mathrm{tr}\ \Big(f(\mu) f(\lambda) f(\nu) f(\lambda)\Big) d\mu d\nu d\lambda \end{aligned}$$

Remark 2.2 Under the null hypothesis of a constant spectral density $f(\lambda) = \Sigma = (\sigma_{i,j})_{i,j=1}^m$ this term simplifies to

$$\tau^2_{M^2,H_0} = 8\pi^2(\text{tr } \Sigma^2)^2.$$

In the special case $m = 1$ we obtain

$$\tau^2_{M^2} = 20\pi \int_{-\pi}^{\pi} f^4(\lambda)d\lambda - 16 \int_{-\pi}^{\pi} f(\lambda)d\lambda \int_{-\pi}^{\pi} f^3(\lambda)d\lambda + \frac{4}{\pi}\Big(\int_{-\pi}^{\pi} f(\lambda)d\lambda\Big)^2 \int_{-\pi}^{\pi} f^2(\lambda)d\lambda$$

in general and $\tau^2_{M^2,H_0} = 8\pi^2\Sigma^4$ under the null hypothesis.

Proof. For the sake of a transparent representation we present the proof in the case $m = 1$ only. The general case follows by exactly the same arguments with an additional amount of notation. In this situation we put $\sigma^2 = \Sigma$ and $\psi_j = c_j^{11}$ for $j \in \mathbb{Z}$, so condition (2.2) can be rewritten as

$$\sum_{j\in\mathbb{Z}} |j||\psi_j| < \infty, \tag{2.6}$$

and by symmetry of both the spectal density function and the periodogram the test statistic in (2.5) reduces to

$$T_n = \frac{1}{2\pi}\Big(T_{n,2} - T^2_{n,1}\Big) = \frac{1}{2\pi}\Big(\frac{2}{n}\sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} I_n(\lambda_k)I_n(\lambda_{k-1}) - \Big(\frac{2}{n}\sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} I_n(\lambda_k)\Big)^2\Big).$$

We will show below that an appropriately standardized version of the vector $(T_{n,1}, T_{n,2})^T$ converges weakly to a normal distribution, that is

$$\sqrt{n}\left(\begin{pmatrix} T_{n,1} \\ T_{n,2} \end{pmatrix} - \begin{pmatrix} \int_{-\pi}^{\pi} f(\lambda)d\lambda \\ 2\pi\int_{-\pi}^{\pi} f^2(\lambda)d\lambda \end{pmatrix}\right) \to N(0, A), \tag{2.7}$$

where the asymptotic covariance matrix is given by

$$A = \begin{pmatrix} 4\pi\int_{-\pi}^{\pi} f^2(\lambda)d\lambda & 16\pi^2\int_{-\pi}^{\pi} f^3(\lambda)d\lambda \\ 16\pi^2\int_{-\pi}^{\pi} f^3(\lambda)d\lambda & 80\pi^3\int_{-\pi}^{\pi} f^4(\lambda)d\lambda \end{pmatrix}.$$

The assertion then follows by a straightforward application of the Delta method to the function $g(x, y) = \frac{1}{2\pi}(y - x^2)$. In order to prove (2.7) we use the approximations

$$\Big|\frac{1}{\sqrt{n}}\sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} I_n(\lambda_k) - \frac{1}{\sqrt{n}}\sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} \tilde{I}_n(\lambda_k)\Big| = o_P(1), \tag{2.8}$$

$$\Big|\frac{1}{\sqrt{n}}\sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} I_n(\lambda_k)I_n(\lambda_{k-1}) - \frac{1}{\sqrt{n}}\sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} \tilde{I}_n(\lambda_k)\tilde{I}_n(\lambda_{k-1})\Big| = o_P(1), \tag{2.9}$$

where each quantity $\tilde{I}_n(\lambda_k)$ is defined by

$$\tilde{I}_n(\lambda_k) \;=\; \left|\psi(e^{-i\lambda_k})\right|^2 \; I_{n,z}(\lambda_k) = \frac{2\pi}{\sigma^2} f(\lambda_k) I_{n,z}(\lambda_k). \tag{2.10}$$

Here, $\psi(z) = \sum_{j\in\mathbb{Z}} \psi_j z^j$ and $I_{n,z}(\lambda)$ denotes the periodogram of the process $\{\mathbf{Z}_t\}_{t\in\mathbb{Z}}$. The proof of (2.8) and (2.9) is complicated and therefore deferred to the appendix in Section 6. From these estimates it follows that it is sufficient to prove assertion (2.7) for the vector

$$\left(\tilde{T}_{n,1}, \tilde{T}_{n,2}\right)^T = \Big(\frac{2}{n}\sum_{k=1}^{\lfloor \frac{n}{2}\rfloor} \tilde{I}_n(\lambda_k), \frac{2}{n}\sum_{k=1}^{\lfloor \frac{n}{2}\rfloor} \tilde{I}_n(\lambda_k)\tilde{I}_n(\lambda_{k-1})\Big)^T.$$

Because $\{\mathbf{Z}_t\}_{t\in\mathbb{Z}}$ is a Gaussian white noise process with variance σ^2, the random variables $\tilde{I}_n(\lambda_k)$ are independent and exponentially distributed for $0 < k < \frac{n}{2}$, that is

$$\tilde{I}_n(\lambda_k) = \frac{2\pi}{\sigma^2} f(\lambda_k)\, I_{n,z}(\lambda_k) \;\sim \exp\left(\frac{1}{2\pi f(\lambda_k)}\right).$$

Therefore a straightforward calculation shows

$$\begin{aligned} E[\tilde{T}_{n,1}] \;&=\; \int_{-\pi}^{\pi} f(\lambda)d\lambda + o(n^{-1/2}), \\ E[\tilde{T}_{n,2}] \;&=\; 2\pi \int_{-\pi}^{\pi} f^2(\lambda)d\lambda + o(n^{-1/2}), \end{aligned}$$

where we have used the Hölder continuity of the spectral density matrix. The assertion now follows by an application of the central limit theorem for m-dependent random variables [see Orey (1958)] and a calculation of the corresponding variances and covariances using the properties of an exponential distribution. Ignoring the boundary terms, we obtain for example for the covariance of the random variables $\tilde{T}_{n,1}$ and $\tilde{T}_{n,2}$

$$\begin{aligned} \lim_{n\to\infty} n\mathrm{Cov}(\tilde{T}_{n,1}, \tilde{T}_{n,2}) \;&=\; \lim_{n\to\infty} n\frac{4}{n^2} \sum_{k_1,k_2=1}^{\lfloor \frac{n}{2}\rfloor} \Big(E(\tilde{I}_n(\lambda_{k_1})\tilde{I}_n(\lambda_{k_2})\tilde{I}_n(\lambda_{k_2-1})) - E(\tilde{I}_n(\lambda_{k_1}))E(\tilde{I}_n(\lambda_{k_2})\tilde{I}_n(\lambda_{k_2-1}))\Big) \\ &=\; \lim_{n\to\infty} \frac{8}{n} \sum_{k=1}^{\lfloor \frac{n}{2}\rfloor} (2\pi) f(\lambda_{k-1})(2\pi)^2 f^2(\lambda_k) = \lim_{n\to\infty} \frac{64\pi^3}{n} \sum_{k=1}^{\lfloor \frac{n}{2}\rfloor} f^3(\lambda_k) \\ &=\; 16\pi^2 \int_{-\pi}^{\pi} f^3(\lambda)d\lambda. \end{aligned}$$

A similar calculation for the variances of $\tilde{T}_{n,1}$ and $\tilde{T}_{n,2}$ yields the assertion of the theorem. $\square$

Theorem 2.1 provides a simple test for the hypotheses (2.3). To be precise, recall that under the null hypothesis $H_0 : f(\lambda) = \Sigma$ the asymptotic variance in Theorem 2.1 simplifies to $\tau^2_{M^2,H_0} = 8\pi^2(\mathrm{tr}\Sigma^2)^2$.

Consequently, if $\hat{\tau}^2_{M^2,H_0}$ is a consistent estimate of $\tau^2_{M^2,H_0}$ it follows from Theorem 2.1 that a consistent asymptotic level α test is obtained by rejecting the null hypothesis if

$$\sqrt{n}\frac{T_n}{\hat{\tau}_{M^2,H_0}} > u_{1-\alpha} \tag{2.11}$$

where $u_{1-\alpha}$ denotes the $(1-\alpha)$ quantile of the standard normal distribution. Observing the representation

$$\Sigma_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\lambda)d\lambda$$

for the parameter corresponding to the best L^2-approximation of the spectral density matrix f by a constant matrix, it follows from the proof of Theorem 2.1 that under the null hypothesis

$$\hat{\Sigma} = \frac{1}{2\pi}T_{n,1} \tag{2.12}$$

converges in probability to Σ. Therefore $\hat{\tau}^2_{M^2,H_0} = 8\pi^2(\text{tr }\hat{\Sigma}^2)^2$ is consistent for $\tau^2_{M^2,H_0}$ as well. The finite sample performance of the corresponding test will be studied in Section 4.

3 Comparing spectral densities

In this section we continue illustrating our approach in a further example comparing the spectral densities of the different components of $\{\mathbf{X}_t\}_{t\in\mathbb{Z}}$. This problem has also found considerable attention in the literature. On one hand, it is closely related to cluster and discriminant analysis [see e.g. Zhang and Taniguchi (1994) or Kakizawa et al. (1998)]. On the other hand, a comparison of the spectral densities can be of own interest [see for example Carmona and Wang (1996), who analyzed Lagrangian velocities of drifters at the surface of the ocean by a comparison of spectra]. Coates and Diggle (1986) compared the spectral densities of two independent time series using periodogram based test statistics and used this method for analyzing wheat price and British gas data. Swanepoel and van Wyk (1986) considered two independent stationary autoregressive processes. Diggle and Fisher (1991) proposed graphical devices to compare periodograms, and a more recent reference is Dette and Paparoditis (2009), who proposed a bootstrap test for the problem of testing for equal spectral densities of m [not necessarily uncorrelated] time series $\{\boldsymbol{X}_{j,t}\}_{t\in\mathbb{Z}}$ $(j=1,\ldots,m)$, that is

$$H_0: \quad f_{11}(\lambda) = \ldots = f_{mm}(\lambda) \tag{3.1}$$

versus

$$H_1: \quad f_{rr}(\lambda) \neq f_{ss}(\lambda) \quad \text{for at least one pair } (r,s), \quad r\neq s.$$

For the construction of an alternative test statistic for the problem (3.1) we investigate the L^2-approximation problem

$$D^2 = \min\left\{\int_{-\pi}^{\pi} \text{tr}((f(\lambda)-g(\lambda))(f(\lambda)-g(\lambda))^*)d\lambda \mid g\in\mathcal{F}_{H_0}\right\},$$

where $\mathcal{F}_{H_0}$ denotes the set of all spectral density matrices $g = (g_{ij})_{i,j=1}^m$ with equal (real) diagonal elements. A similar calculation as in the previous section yields for $g \in \mathcal{F}_{H_0}$

$$
\begin{aligned}
(3.2)\quad & \int_{-\pi}^{\pi} \operatorname{tr}((f(\lambda)-g(\lambda))(f(\lambda)-g(\lambda))^* d\lambda = \sum_{i,j=1}^m \int_{-\pi}^{\pi} |f_{ij}(\lambda)-g_{ij}(\lambda)|^2 d\lambda \\
& \geq \sum_{i=1}^m \int_{-\pi}^{\pi} |f_{ii}(\lambda)-g_{11}(\lambda)|^2 d\lambda = \sum_{i=1}^m \int_{-\pi}^{\pi} |f_{ii}(\lambda)-h(\lambda)+h(\lambda)-g_{11}(\lambda)|^2 d\lambda \\
& \geq \sum_{i=1}^m \int_{-\pi}^{\pi} |f_{ii}(\lambda)-h(\lambda)|^2 d\lambda + 2\sum_{i=1}^m \int_{-\pi}^{\pi} (f_{ii}(\lambda)-h(\lambda))(h(\lambda)-g_{11}(\lambda)) d\lambda \\
& = \sum_{i=1}^m \int_{-\pi}^{\pi} |f_{ii}(\lambda)-h(\lambda)|^2 d\lambda = \frac{1}{m}\Big\{(m-1)\sum_{i=1}^m \int_{-\pi}^{\pi} f_{ii}^2(\lambda)d\lambda - 2\sum_{1\leq i<j\leq m}\int_{-\pi}^{\pi} f_{ii}(\lambda)f_{jj}(\lambda)d\lambda\Big\}
\end{aligned}
$$

with $h(\lambda) = \frac{1}{m}\sum_{i=1}^m f_{ii}(\lambda)$, and there is equality in (3.2) for $g(\lambda) = ((1-\delta_{ij})f_{ij}(\lambda) + \delta_{ij}h(\lambda))_{i,j=1}^m$ [here δ_{ij} denotes the Kronecker symbol]. Therefore we consider the L^2 distance

$$
\begin{aligned}
D^2 &= \sum_{1\leq i<j\leq m} \int_{-\pi}^{\pi} (f_{ii}(\lambda)-f_{jj}(\lambda))^2 d\lambda \\
&= (m-1)\sum_{i=1}^m \int_{-\pi}^{\pi} f_{ii}^2(\lambda)d\lambda - 2\sum_{1\leq i<j\leq m}\int_{-\pi}^{\pi} f_{ii}(\lambda)f_{jj}(\lambda)d\lambda,
\end{aligned}
$$

which obviously vanishes if and only if the null hypothesis is satisfied. In order estimate the quantity D^2 we define

$$
(3.3)\qquad T_n^{(ij)} = \frac{2}{n}\sum_{k=1}^{\lfloor \frac{n}{2}\rfloor} I_{n,ii}(\lambda_k) I_{n,jj}(\lambda_{k-1}); \qquad 1\leq i\leq j\leq m,
$$

where $I_{n,ii}$ denotes the ith diagonal element of the periodogram I_n defined in (2.4) for any $i = 1,\ldots,m$. We consider the test statistic

$$
(3.4)\qquad D_n = \frac{m-1}{2\pi}\sum_{i=1}^m T_n^{(ii)} - \frac{1}{\pi}\sum_{1\leq i<j\leq m} T_n^{(ij)}.
$$

The following result specifies the asymptotic distribution of the statistic D_n under the null hypothesis and alternative.

Theorem 3.1 *If $\{\boldsymbol{X}_t\}_{t\in\mathbb{Z}}$ is a stationary process satisfying assumptions (2.1) and (2.2) with Hölder continuous spectral density matrix of order $L > 1/2$, then we have for the statistic D_n from (3.4)*

$$
\sqrt{n}\Big(D_n - D^2\Big) \to N(0, \tau_{D^2}^2),
$$

where the asymptotic variance is given by

$$
\begin{aligned}
\tau_{D^2}^2 \;=\; & \pi\Big(4(m-1)^2 \sum_{1\le i,j\le m} \int_{-\pi}^{\pi} \Big\{ f_{ij}^2(\lambda) f_{ji}^2(\lambda) + 4 f_{ii}(\lambda) f_{ij}(\lambda) f_{ji}(\lambda) f_{jj}(\lambda) \Big\} d\lambda \\
& -16(m-1) \sum_{k=1}^{m} \sum_{1\le i<j\le m} \int_{-\pi}^{\pi} \Big\{ 2 f_{kk}(\lambda) f_{ki}(\lambda) f_{ik}(\lambda) f_{jj}(\lambda) + 2 f_{kk}(\lambda) f_{kj}(\lambda) f_{jk}(\lambda) f_{ii}(\lambda) \\
& + f_{ki}(\lambda) f_{ik}(\lambda) f_{kj}(\lambda) f_{jk}(\lambda) \Big\} d\lambda \\
& +16 \sum_{1\le k<l\le m} \sum_{1\le i<j\le m} \int_{-\pi}^{\pi} \Big\{ 2 f_{kk}(\lambda) f_{ll}(\lambda) f_{ij}(\lambda) f_{ji}(\lambda) + f_{kk}(\lambda) f_{ii}(\lambda) f_{lj}(\lambda) f_{jl}(\lambda) \\
& + f_{ll}(\lambda) f_{jj}(\lambda) f_{ik}(\lambda) f_{ki}(\lambda) + f_{ki}(\lambda) f_{ik}(\lambda) f_{lj}(\lambda) f_{jl}(\lambda) \Big\} d\lambda \Big).
\end{aligned}
$$

Remark 3.2 If the null hypothesis (3.1) is satisfied, we have

$$\sqrt{n} D_n \to N(0, \tau_{D^2,H_0}^2),$$

and the asymptotic variance τ_{D^2,H_0}^2 is given by

$$
\begin{aligned}
\tau_{D^2,H_0}^2 \;=\; & \pi\Big(\frac{4m(m-1)(11-m)}{3} \int_{-\pi}^{\pi} f_{11}^4(\lambda) d\lambda + 8(m-1)^2 \sum_{1\le i<j\le m} \int_{-\pi}^{\pi} f_{ij}^2(\lambda) f_{ji}^2(\lambda) d\lambda \\
& + \sum_{1\le i<j\le m} \Big(32(j(i-1) + (m-i)(m-j)) + 16(m-1)(3m-8) \Big) \int_{-\pi}^{\pi} f_{11}^2(\lambda) f_{ij}(\lambda) f_{ji}(\lambda) d\lambda \\
& +32 \sum_{1\le i<j\le m} \sum_{k=i+1}^{m} \sum_{l=j+1}^{m} \int_{-\pi}^{\pi} f_{ik}(\lambda) f_{ki}(\lambda) f_{jl}(\lambda) f_{lj}(\lambda) d\lambda \\
& -16(m-1) \sum_{1\le i\ne j\ne k\ne i\le m} \int_{-\pi}^{\pi} \Big\{ 2 f_{kk}(\lambda) f_{ki}(\lambda) f_{ik}(\lambda) f_{jj}(\lambda) + 2 f_{kk}(\lambda) f_{kj}(\lambda) f_{jk}(\lambda) f_{ii}(\lambda) \\
& + f_{ki}(\lambda) f_{ik}(\lambda) f_{kj}(\lambda) f_{jk}(\lambda) \Big\} d\lambda.
\end{aligned}
$$

Remark 3.3 In the case of comparing the spectral densities of two samples ($m = 2$) the asymptotic variance in Theorem 3.1 becomes

$$
\begin{aligned}
(3.5)\quad \tau_{D^2}^2 \;=\; & 20\pi \int_{-\pi}^{\pi} f_{11}^4(\lambda) d\lambda - 32\pi \int_{-\pi}^{\pi} f_{11}^3(\lambda) f_{22}(\lambda) d\lambda - 48\pi \int_{-\pi}^{\pi} f_{11}^2(\lambda) f_{12}(\lambda) f_{21}(\lambda) d\lambda \\
& +48\pi \int_{-\pi}^{\pi} f_{11}^2(\lambda) f_{22}^2(\lambda) d\lambda + 64\pi \int_{-\pi}^{\pi} f_{11}(\lambda) f_{12}(\lambda) f_{21}(\lambda) f_{22}(\lambda) d\lambda + 8\pi \int_{-\pi}^{\pi} f_{12}^2(\lambda) f_{21}^2(\lambda) d\lambda \\
& -48\pi \int_{-\pi}^{\pi} f_{22}^2(\lambda) f_{12}(\lambda) f_{21}(\lambda) d\lambda - 32\pi \int_{-\pi}^{\pi} f_{11}(\lambda) f_{22}^3(\lambda) d\lambda + 20\pi \int_{-\pi}^{\pi} f_{22}^4(\lambda) d\lambda.
\end{aligned}
$$

which yields under the null hypothesis

$$\tau^2_{D^2,H_0} = 24\pi \int_{-\pi}^{\pi} f_{11}^4(\lambda)d\lambda - 32\pi \int_{-\pi}^{\pi} f_{11}^2(\lambda)f_{12}(\lambda)f_{21}(\lambda)d\lambda + 8\pi \int_{-\pi}^{\pi} f_{12}^2(\lambda)f_{21}^2(\lambda)d\lambda. \tag{3.6}$$

A consistent estimator of this quantity is given by

$$\begin{aligned} \hat{\tau}^2_{D^2,H_0} &= \frac{6}{n\pi^2} \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} I_{n,11}(\lambda_k) I_{n,11}(\lambda_{k-1}) I_{n,11}(\lambda_{k-2}) I_{n,11}(\lambda_{k-3}) \\ &\quad - \frac{8}{n\pi^2} \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} I_{n,11}(\lambda_k) I_{n,11}(\lambda_{k-1}) I_{n,12}(\lambda_{k-2}) I_{n,21}(\lambda_{k-3}) \\ &\quad + \frac{2}{n\pi^2} \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} I_{n,12}(\lambda_k) I_{n,12}(\lambda_{k-1}) I_{n,21}(\lambda_{k-2}) I_{n,21}(\lambda_{k-3}), \end{aligned} \tag{3.7}$$

and similar estimates can be derived for the variance under the null hypothesis in Theorem 3.1 in the case $m > 2$. Thus the null hypothesis (3.1) is rejected whenever

$$\sqrt{n} \frac{D_n}{\hat{\tau}_{D^2,H_0}} > u_{1-\alpha}, \tag{3.8}$$

where $u_{1-\alpha}$ denotes the $(1-\alpha)$ quantile of the standard normal distribution.

Proof of Theorem 3.1. For the sake of brevity we restrict ourselves to proving the result in the case of two samples. The general assertion follows by exactly the same arguments with an additional amount of notation. For the case $m = 2$ we recall the definition of the estimates in (3.3) and show the weak convergence of the vector $(T_n^{(11)}, T_n^{(12)}, T_n^{(22)})^T$, that is

$$\sqrt{n} \left[\begin{pmatrix} T_n^{(11)} \\ T_n^{(12)} \\ T_n^{(22)} \end{pmatrix} - \begin{pmatrix} 2\pi \int_{-\pi}^{\pi} f_{11}^2(\lambda)d\lambda \\ 2\pi \int_{-\pi}^{\pi} f_{11}(\lambda)f_{22}(\lambda)d\lambda \\ 2\pi \int_{-\pi}^{\pi} f_{22}^2(\lambda)d\lambda \end{pmatrix} \right] \to N(0, \Lambda) \tag{3.9}$$

where the elements of the (symmetric) matrix $\Lambda = (\lambda_{ij})_{i,j=1}^3$ are given by

$$\begin{aligned} \lambda_{11} &= 80\pi^3 \int_{-\pi}^{\pi} f_{11}^4(\lambda)d\lambda, \\ \lambda_{12} &= 16\pi^3 \Big(2 \int_{-\pi}^{\pi} f_{11}^3(\lambda)f_{22}(\lambda)d\lambda + 3 \int_{-\pi}^{\pi} f_{11}^2(\lambda)f_{12}(\lambda)f_{21}(\lambda)d\lambda \Big), \\ \lambda_{13} &= 16\pi^3 \Big(\int_{-\pi}^{\pi} f_{12}^2(\lambda)f_{21}^2(\lambda)d\lambda + 4 \int_{-\pi}^{\pi} f_{11}(\lambda)f_{12}(\lambda)f_{21}(\lambda)f_{22}(\lambda)d\lambda, \\ \lambda_{22} &= 16\pi^3 \Big(3 \int_{-\pi}^{\pi} f_{11}^2(\lambda)f_{22}^2(\lambda)d\lambda + 2 \int_{-\pi}^{\pi} f_{11}(\lambda)f_{12}(\lambda)f_{21}(\lambda)f_{22}(\lambda)d\lambda \Big), \\ \lambda_{23} &= 16\pi^3 \Big(2 \int_{-\pi}^{\pi} f_{11}(\lambda)f_{22}^3(\lambda)d\lambda + 3 \int_{-\pi}^{\pi} f_{22}^2(\lambda)f_{12}(\lambda)f_{21}(\lambda)d\lambda \Big), \\ \lambda_{33} &= 80\pi^3 \int_{-\pi}^{\pi} f_{22}^4(\lambda)d\lambda. \end{aligned}$$

Theorem 3.1 then follows again by the Delta method, observing that

$$D_n = g(T_{n,2}^{(11)}, T_{n,1}^{(12)}, T_{n,2}^{(22)}) = \frac{1}{2\pi} T_{n,1}^{(11)} - \frac{1}{\pi} T_{n,1}^{(12)} + \frac{1}{2\pi} T_{n,2}^{(22)},$$

where the function g is defined by $g(x,y,z) = \frac{1}{2\pi}x - \frac{1}{\pi}y + \frac{1}{2\pi}z$. In order to prove (3.9), we note that a similar reasoning as for (2.8) and (2.9) (see Section 6) yields the estimate

$$(3.10) \qquad \Big| \frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} I_{n,ii}(\lambda_k) I_{n,jj}(\lambda_{k-1}) - \frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} \tilde{I}_{n,ii}(\lambda_k) \tilde{I}_{n,jj}(\lambda_{k-1}) \Big| = o_P(1)$$

for arbitrary $1 \leq i, j \leq 2$, where $\tilde{I}_{n,ij}(\lambda_k)$ denotes the element in the position (i,j) of the matrix

$$\tilde{I}_n(\lambda_k) = 2\pi f^{1/2}(\lambda_k) \Sigma^{-1/2} I_{n,z}(\lambda_k) \Sigma^{-1/2} (f^{1/2}(\lambda_k))^*.$$

The claim in (3.9) thus follows, if a corresponding statement for the vector $(\tilde{T}_n^{(11)}, \tilde{T}_n^{(12)}, \tilde{T}_n^{(22)})^T$ can be established, where the statistics $\tilde{T}_n^{(ij)}$ are defined by

$$\tilde{T}_n^{(ij)} = \frac{2}{n} \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} \tilde{I}_{n,ii}(\lambda_k) \tilde{I}_{n,jj}(\lambda_{k-1}).$$

The assertions now are a consequence of the central limit theorem in Orey (1958) [note that the terms in the sum are 1-dependent], where the elements in the covariance are obtained by a careful calculation observing

$$E\Big(\tilde{I}_{n,ij}(\lambda_k) \tilde{I}_{n,rs}(\lambda_k)\Big) \;=\; (2\pi)^2 \Big(f_{ij}(\lambda_k) f_{rs}(\lambda_k) + f_{is}(\lambda_k) f_{rj}(\lambda_k) \Big)$$

for $\lambda_k \in (0,\pi)$ and any $1 \leq i,j,r,s \leq 2$ [see Hannan (1970)]. □

4 Finite sample properties

In this section we will present a small simulation study to investigate the finite sample properties of the proposed test statistic. We will consider the problems of testing for a constant spectral density and comparing the spectral densities of two time series separately. All presented results are based on 1000 simulation runs.

Example 4.1: Testing for a constant spectral density. In order to investigate the testing problem (2.3) we consider the models

$$(4.1) \qquad X_t \;=\; Z_t$$

$$(4.2) \qquad X_t \;=\; Z_t + \frac{1}{5} Z_{t+1},$$

corresponding to null hypothesis and alternative, respectively. Here $\{Z_t\}_{t\in\mathbb{Z}}$ is a Gaussian white noise process with variance $\sigma^2 = 1$. Note that for model (4.2) the spectral density is given by

$f(\lambda) = \frac{1}{2\pi}\left(\frac{26}{25} + \frac{2}{5}\cos(\lambda)\right)$. In Table 1 we show the rejection probabilities of the test (2.11) for various sample sizes, where the asymptotic variance has been estimated by (2.12). We observe a rather accurate approximation of the nominal level and reasonable rejection probabilities under the alternative. Simulations of other scenarios showed a similar picture and are not displayed for the sake of brevity.

	(4.1)			(4.2)		
	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 15\%$	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 15\%$
$n = 128$	0.073	0.098	0.125	0.249	0.296	0.338
$n = 256$	0.066	0.105	0.125	0.289	0.359	0.422
$n = 512$	0.062	0.093	0.139	0.389	0.487	0.557
$n = 1024$	0.056	0.099	0.141	0.569	0.650	0.709

Table 1: *Rejection probabilities of the test (2.11) for the hypothesis of a constant spectral density under the null hypothesis and alternative.*

Example 4.2: Comparing spectral densities of stationary time series. In this example we study the testing problem (3.1) in the case $m = 2$, where the stationary time series is given by $\{(X_{1,t}, X_{2,t})\}_{t\in\mathbb{Z}}$ with

$$\begin{aligned} X_{1,t} &= Z_{1,t} - \beta_1 Z_{1,t-1} - \beta_2 Z_{1,t-2} \\ X_{2,t} &= Z_{2,t} - \beta_1 Z_{2,t-1} \end{aligned}$$

Here $\beta_1 = 0.8$, and $\{(Z_{1,t}, Z_{2,t})^T\}_{t\in\mathbb{Z}}$ is an independent centered stationary Gaussian process with covariance matrix

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

and the choice β_2 corresponds to either the null hypothesis of equal spectral densities ($\beta_2 = 0$) or to the alternative ($\beta_2 \neq 0$). In Table 2 we present the rejection probabilities of the test (3.8) for various sample sizes where we used the variance estimate (3.7). We observe again that the nominal level is well approximated and that the alternatives are clearly detected in all cases. It is interesting to note that the power of the test is increasing with the correlation ρ of the Gaussian process. A similar observation was also made by Dette and Paparoditis (2009) for a test based on a kernel estimate of the spectral density matrix.

In the present situation the empirical observations can be explained by the asymptotic theory presented in Section 3. First, note that under the alternative the probability of rejecting the null hypothesis of equal spectral densities by the test (3.8) is approximately given by

$$P\Big(\sqrt{n}\frac{D_n}{\hat{\tau}_{D,H_0}} > u_{1-\alpha}\Big) \approx \Phi\Big(\sqrt{n}\frac{D^2}{\tau_{D^2}} - u_{1-\alpha}\frac{\tau_{D^2,H_0}}{\tau_{D^2}}\Big) \approx \Phi\Big(\sqrt{n}\frac{D^2}{\tau_{D^2}}\Big) \tag{4.3}$$

	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.8$		
$\beta_2 = 0$	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 15\%$	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 15\%$	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 15\%$
$n = 128$	0.074	0.112	0.153	0.078	0.125	0.172	0.054	0.084	0.121
$n = 256$	0.057	0.115	0.159	0.053	0.096	0.153	0.052	0.096	0.140
$n = 512$	0.046	0.095	0.143	0.044	0.090	0.128	0.042	0.083	0.127
$n = 1024$	0.059	0.109	0.156	0.049	0.099	0.146	0.043	0.085	0.131
$\beta_2 = 0.5$	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 15\%$	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 15\%$	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 15\%$
$n = 128$	0.155	0.270	0.355	0.202	0.297	0.382	0.194	0.352	0.455
$n = 256$	0.216	0.347	0.431	0.231	0.366	0.452	0.343	0.497	0.613
$n = 512$	0.293	0.455	0.552	0.379	0.543	0.635	0.568	0.712	0.813
$n = 1024$	0.477	0.638	0.737	0.556	0.706	0.794	0.824	0.908	0.942

Table 2: *Simulated rejection probabilities of the test (3.8) for the hypothesis (3.1) of equal spectral densities.*

where Φ denotes the distribution function of the standard normal distribution and τ^2_{D,H_0} and τ^2_D are the asymptotic variances under the null hypothesis and alternative, respectively [see equations (3.5) and (3.6)]. This shows that the power is increasing with the ratio D^2/τ_{D^2}. In Figure 1 we display the ratio

$$\rho \to p(\rho) = \frac{D^2(\rho)}{\tau_D(\rho)} \tag{4.4}$$

as a function of $\rho \in [-1, 1]$ and we observe that the asymptotic power (4.3) is an increasing function of $|\rho|$. This confirms our empirical observations in Table 2.

5 Conclusions

In this paper we have illustrated an alternative concept for constructing tests for nonparametric hypotheses in stationary time series with a linear representation of the form (2.1). Our approach is based on an estimate of the L^2-distance between the spectral density matrix and its best approximation under the null hypothesis and does not require the specification of a smoothing parameter. The test statistic is constructed from simple estimates of integrated components of the spectral density matrix and follows an asymptotic normal distribution. For the sake of a clear presentation and brevity we have restricted ourselves to the problem of testing for a constant spectral density matrix and to the comparison of the spectral densities of several correlated time series.

We conclude this paper with a few remarks on generalizations of our approach. First, the generalization of the approach to other hypotheses is obvious, if we consider the minimal distance

$$M^2 = \min \Big\{ \int_{-\pi}^{\pi} \mathrm{tr}\{(f(\lambda) - g(\lambda))(f(\lambda) - g(\lambda))^*\} d\lambda \mid g \in \mathcal{F}_{H_0} \Big\}, \tag{5.1}$$

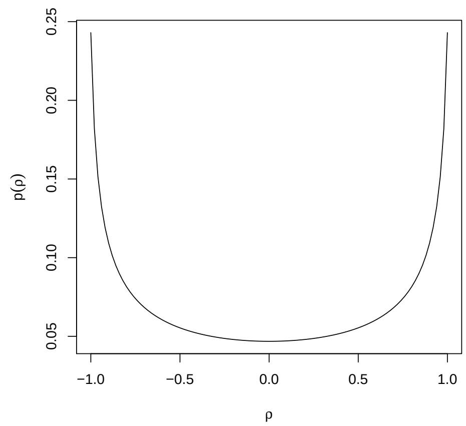

Figure 1: *The function $p(\rho)$ defined in (4.4) for various values of ρ*

where $\mathcal{F}_{H_0}$ denotes the set of all spectral densities satisfying the null hypothesis under consideration. In most cases this minimal distance can be given explicitly in terms of integrals of elements of the spectral density matrix over the full frequency domain $[-\pi, \pi]$. Typical examples include the problem of no correlation and separability. The first problem corresponds to the hypothesis

$$H_0 : f_{ij}(\lambda) = 0, \quad 1 \ \leq i < j \leq m$$

in the frequency domain. In this case, it is easy to see that the minimum in (5.1) is given by

$$M^2 = 2 \sum_{1 \leq i < j \leq m} \int_{-\pi}^{\pi} |f_{ij}(\lambda)|^2 d\lambda$$

which can be estimated easily using the approach described in Sections 2 and 3. The (appropriately standardized) test statistic is asymptotically normal distributed and it only remains to estimate the asymptotic variance under the null hypothesis. The second example corresponds to the hypothesis of separability

$$H_0 : f(\lambda) = \Sigma f_0(\lambda)$$

[see e.g. Matsuda and Yajima (2004)], where Σ denotes a positive definite matrix and $f_0(\lambda)$ is a real-valued function which is integrable on $[-\pi, \pi]$. Without loss of generality one can take Σ to be the variance of the process $\{\boldsymbol{X}_t\}_{t \in \mathbb{Z}}$, which gives

$$f_0(\lambda) = \frac{1}{m} \mathrm{tr}\Big(f(\lambda) V_\Sigma^{-1}\Big) = \frac{1}{m} \sum_{i=1}^{m} \frac{f_{ii}(\lambda)}{\int_{-\pi}^{\pi} f_{ii}(\lambda) d\lambda}$$

with $V_\Sigma = \mathrm{diag}(\sigma_1^2, \dots, \sigma_m^2)$ being the diagonal matrix whose entries are the variances of $X_{t,i}$ [see also Eichler (2008)]. Using the same approach as in Section 2 it can be shown that the distance in (5.1) becomes minimal for

$$\Sigma_0 = \frac{\int_{-\pi}^{\pi} f(\lambda) f_0(\lambda) d\lambda}{\int_{-\pi}^{\pi} f_0^2(\lambda) d\lambda},$$

which gives

$$M^2 = \int_{-\pi}^{\pi} \operatorname{tr}(f(\lambda) f^*(\lambda)) d\lambda - \frac{\operatorname{tr}((\int_{-\pi}^{\pi} f(\lambda) f_0(\lambda) d\lambda)(\int_{-\pi}^{\pi} f(\lambda) f_0(\lambda) d\lambda)^*)}{\int_{-\pi}^{\pi} f_0^2(\lambda) d\lambda}.$$

This quantity can also be estimated easily by the methods described in the previous paragraphs. Other hypotheses can be treated similarly, and the concept presented here is applicable as soon as the minimal distance can be represented as a functional of the spectral density matrix.

Second, the restriction to Gaussian innovations was made throughout this paper in order to keep the arguments simple. In fact, similar results can be derived for arbitrary linear processes of the form (2.1), as long as the sequence of innovations $\{\mathbf{Z}_t\}_{t \in \mathbb{Z}}$ is independent identically distributed with $E[\mathbf{Z}_t] = 0$ and existing moments of eighth order. Note however that the simple form of the variances $\tau_{M^2}^2$ and $\tau_{D^2}^2$ in both theorems is due to special properties of the normal distribution and that cumulants of higher orders show up in general. For example, if we are in the situation of Theorem 2.1 (and in the one-dimensional case $m = 1$), we obtain a central limit theorem $\sqrt{n}(T_n - M^2) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \overline{\tau}_{M^2}^2)$ with variance

$$\begin{aligned} \overline{\tau}_{M^2}^2 \; = \; & 20\pi \int_{-\pi}^{\pi} f^4(\lambda) d\lambda - 16 \int_{-\pi}^{\pi} f(\lambda) d\lambda \int_{-\pi}^{\pi} f^3(\lambda) d\lambda + \frac{4}{\pi} \Big(\int_{-\pi}^{\pi} f(\lambda) d\lambda \Big)^2 \int_{-\pi}^{\pi} f^2(\lambda) d\lambda \\ & + q \Big(\frac{1}{\pi} \Big(\int_{-\pi}^{\pi} f(\lambda) d\lambda \Big)^2 - 2 \int_{-\pi}^{\pi} f^2(\lambda) d\lambda \Big)^2, \end{aligned} \tag{5.2}$$

where $q = \kappa_4/\sigma^4$ and $\kappa_4 = E[Z_t^4] - 3\sigma^2$ denotes the fourth cumulant of Z_t. This term coincides with $\tau_{M^2}^2$ if Z_t is normally distributed, as we have $\kappa_4 = 0$ in this case. Moreover, under the null hypothesis of a constant spectral density the asymptotic variance in (5.2) does also not depend on the fourth cumulant. A similar phenomenon can be observed for the tests proposed by Eichler (2008) for a general class of hypotheses [see Dette and Hildebrandt (2010)].

6 Appendix: some technical details

In this section, we show the estimates (2.8) and (2.9) which are the main ingredients for the proof of both Theorem 2.1 and 3.1. To this end, let $R_n(\lambda_k) = I_n(\lambda_k) - \tilde{I}_n(\lambda_k)$ denote the differences in (2.8), which by a standard argument [see Brockwell and Davis (1991), p. 347] can be represented as

$$\begin{aligned} R_n(\lambda_k) \; &= \; \psi(e^{-i\lambda_k}) J_{n,z}(\lambda_k) \overline{Y_n(\lambda_k)} + \overline{\psi(e^{-i\lambda_k})} \; \overline{J_{n,z}(\lambda_k)} Y_n(\lambda_k) + |Y_n(\lambda_k)|^2 \\ &= \; R_{n,1}(\lambda_k) + R_{n,2}(\lambda_k) + |Y_n(\lambda_k)|^2, \end{aligned} \tag{6.1}$$

where we have used the notation

$$
\begin{aligned}
\psi(e^{-i\lambda}) &= \sum_{j=-\infty}^{\infty} \psi_j e^{-i\lambda j} \\
U_{n,j}(\lambda) &= \sum_{t=1-j}^{n-j} Z_t e^{-i\lambda t} - \sum_{t=1}^{n} Z_t e^{-i\lambda t} \\
Y_n(\lambda) &= \frac{1}{\sqrt{n}} \sum_{j=-\infty}^{\infty} \psi_j e^{-i\lambda j} U_{n,j}(\lambda) \qquad (6.2)
\end{aligned}
$$

and the quantities $R_{n,1}$ and $R_{n,2}$ are defined in an obvious way. Using the symmetry of the periodogram [and ignoring boundary effects] it is our aim to show

$$
\Big| \frac{1}{\sqrt{n}} \sum_{k=-\lfloor \frac{n-1}{2} \rfloor}^{\lfloor \frac{n}{2} \rfloor} R_n(\lambda_k) \Big| = o_P(1) \qquad (6.3)
$$

in order to obtain (2.8). It is well known [see Brockwell and Davis (1991)] that

$$
E|Y_n(\lambda_k)|^2 \leq E(|Y_n(\lambda_k)|^4)^{\frac{1}{2}} = O(n^{-1}), \qquad (6.4)
$$

which yields the assertion for the third term in (6.1). The remaining two terms can be treated similarly, so for the sake of brevity only $R_{n,1}(\lambda_k)$ is considered here. This quantity can be decomposed as follows: First we set

$$
\begin{aligned}
Y_n(\lambda_k) &= \frac{1}{\sqrt{n}} \sum_{l=-\infty}^{-1} \psi_l e^{-i\lambda_k l} U_{n,l}(\lambda_k) + \frac{1}{\sqrt{n}} \sum_{l=0}^{\infty} \psi_l e^{-i\lambda_k l} U_{n,l}(\lambda_k) \qquad (6.5) \\
&= H_n^-(\lambda_k) + H_n^+(\lambda_k),
\end{aligned}
$$

where the last identity defines the expressions H_n^- and H_n^+. By definition we have $\lambda_k = \frac{2\pi k}{n}$, so a straightforward calculation yields the representation

$$
\begin{aligned}
H_n^+(\lambda_k) &= \frac{1}{\sqrt{n}} \sum_{l=0}^{\infty} \sum_{r=1}^{l} \psi_l (Z_{r-l} - Z_{n+r-l}) e^{-i\lambda_k r}, \qquad (6.6) \\
H_n^-(\lambda_k) &= \frac{1}{\sqrt{n}} \sum_{l=-\infty}^{-1} \sum_{r=1+l}^{0} \psi_l (Z_{n+r-l} - Z_{r-l}) e^{-i\lambda_k r}.
\end{aligned}
$$

Setting

$$
\begin{aligned}
R_{n,1}^+ &= \sum_{k=-\lfloor \frac{n-1}{2} \rfloor}^{\lfloor \frac{n}{2} \rfloor} \psi(e^{-i\lambda_k}) J_{n,z}(\lambda_k) \overline{H_n^+(\lambda_k)} \\
&= \frac{1}{n} \sum_{k=-\lfloor \frac{n-1}{2} \rfloor}^{\lfloor \frac{n}{2} \rfloor} \sum_{j=-\infty}^{\infty} \sum_{l=0}^{\infty} \sum_{r=1}^{l} \sum_{t=1}^{n} \psi_j \psi_l Z_t (Z_{r-l} - Z_{n+r-l}) e^{-i\lambda_k (j+t-r)}
\end{aligned}
$$

and similarly

$$R_{n,1}^- = \sum_{k=-\lfloor \frac{n-1}{2} \rfloor}^{\lfloor \frac{n}{2} \rfloor} \psi(e^{-i\lambda_k}) J_{n,z}(\lambda_k) \overline{H_n^-(\lambda_k)}$$

we are left to show that both $n^{-1/2}E|R_{n,1}^+|$ and $n^{-1/2}E|R_{n,1}^-|$ converge to zero, and we restrict ourselves to the proof of the first claim. Note for any fixed two integers j and r that the relation

$$\sum_{k=-\lfloor \frac{n-1}{2} \rfloor}^{\lfloor \frac{n}{2} \rfloor} e^{-i\lambda_k(j+t-r)} = \begin{cases} n, & t = r - j \bmod n \\ 0, & \text{otherwise} \end{cases} \tag{6.7}$$

holds, from which we conclude

$$R_{n,1}^+ = \sum_{j=-\infty}^{\infty} \sum_{l=0}^{\infty} \sum_{r=1}^{l} \psi_j \psi_l Z_{t^*}(Z_{r-l} - Z_{n+r-l}),$$

where $t^* = t(r,j)$ is the unique $t \in \{1, \dots, n\}$, such that $t = r - j \bmod n$ holds. Therefore

$$E|R_{n,1}^+| \le 2\sigma^2 \sum_{j=-\infty}^{\infty} \sum_{l=0}^{\infty} \sum_{r=1}^{l} |\psi_j||\psi_l| \le 2\sigma^2 \sum_{j=-\infty}^{\infty} |\psi_j| \sum_{l=-\infty}^{\infty} |l|\,|\psi_l| = O(1),$$

where we have used condition (2.6) for the last estimate. Analogously, the estimate $E|R_{n,1}^-| = O(1)$ follows, and a similar argument for the sum involving $R_{n,2}(\lambda_k)$ yields assertion (2.8).
We now turn to the proof of estimate (2.9). From (6.3), the Cauchy-Schwarz inequality and the symmetry of the periodogram we have

$$\begin{aligned} & E\Big|\frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} \Big(I_n(\lambda_k) I_n(\lambda_{k-1}) - \tilde{I}_n(\lambda_k) \tilde{I}_n(\lambda_{k-1}) \Big) \Big| \\ = \; & E\Big|\frac{1}{2\sqrt{n}} \sum_{k=-\lfloor \frac{n-1}{2} \rfloor}^{\lfloor \frac{n}{2} \rfloor} \Big(I_n(\lambda_k) I_n(\lambda_{k-1}) - \tilde{I}_n(\lambda_k) \tilde{I}_n(\lambda_{k-1}) \Big) \Big| + o(1) \\ = \; & E\Big|\frac{1}{2\sqrt{n}} \sum_{k=-\lfloor \frac{n-1}{2} \rfloor}^{\lfloor \frac{n}{2} \rfloor} \tilde{I}_n(\lambda_k) R_n(\lambda_{k-1}) + \frac{1}{2\sqrt{n}} \sum_{k=-\lfloor \frac{n-1}{2} \rfloor}^{\lfloor \frac{n}{2} \rfloor} \tilde{I}_n(\lambda_{k-1}) R_n(\lambda_k) \Big| + o(1). \end{aligned} \tag{6.8}$$

Once we have shown the claim

$$E\Big(\frac{1}{\sqrt{n}} \sum_{k=-\lfloor \frac{n-1}{2} \rfloor}^{\lfloor \frac{n}{2} \rfloor} \tilde{I}_n(\lambda_k) R_n(\lambda_{k-1}) \Big)^2 = o(1) \tag{6.9}$$

and an analogous estimate with λ_k replaced by λ_{k-1} and vice versa, (2.9) follows.

A proof of (6.9) is quite delicate. Note first from a simple calculation that

$$E|\tilde{I}_n(\lambda_k)|^p < C_p < \infty \tag{6.10}$$

uniformly in k, for any $p > 0$ and some generic constants $C_p > 0$. Now recall the decomposition (6.1) for the remainder term $R_n(\lambda_k)$. We have

$$E\Big(\frac{1}{\sqrt{n}}\sum_{k=-\lfloor\frac{n-1}{2}\rfloor}^{\lfloor\frac{n}{2}\rfloor}\tilde{I}_n(\lambda_k)R_n(\lambda_{k-1})\Big)^2 = \frac{1}{n}\sum_{j=-\lfloor\frac{n-1}{2}\rfloor}^{\lfloor\frac{n}{2}\rfloor}\sum_{k=-\lfloor\frac{n-1}{2}\rfloor}^{\lfloor\frac{n}{2}\rfloor}E\Big(\tilde{I}_n(\lambda_j)R_n(\lambda_{j-1})\tilde{I}_n(\lambda_k)R_n(\lambda_{k-1})\Big),$$

and therefore we require very accurate estimates for the terms in the decomposition

$$\begin{aligned}
R_n(\lambda_{j-1})R_n(\lambda_{k-1}) \;=\; & \psi(e^{-i\lambda_{j-1}})J_{n,z}(\lambda_{j-1})Y_n(-\lambda_{j-1})\psi(e^{-i\lambda_{k-1}})J_{n,z}(\lambda_{k-1})Y_n(-\lambda_{k-1})\\
&+\psi(e^{-i\lambda_{j-1}})J_{n,z}(\lambda_{j-1})Y_n(-\lambda_{j-1})\psi(e^{i\lambda_{k-1}})J_{n,z}(-\lambda_{k-1})Y_n(\lambda_{k-1})\\
&+\psi(e^{i\lambda_{j-1}})J_{n,z}(-\lambda_{j-1})Y_n(\lambda_{j-1})\psi(e^{-i\lambda_{k-1}})J_{n,z}(\lambda_{k-1})Y_n(-\lambda_{k-1})\\
&+\psi(e^{i\lambda_{j-1}})J_{n,z}(-\lambda_{j-1})Y_n(\lambda_{j-1})\psi(e^{i\lambda_{k-1}})J_{n,z}(-\lambda_{k-1})Y_n(\lambda_{k-1})\\
&+\psi(e^{-i\lambda_{j-1}})J_{n,z}(\lambda_{j-1})Y_n(-\lambda_{j-1})|Y_n(\lambda_{k-1})|^2\\
&+\psi(e^{i\lambda_{j-1}})J_{n,z}(-\lambda_{j-1})Y_n(\lambda_{j-1})|Y_n(\lambda_{k-1})|^2\\
&+\psi(e^{-i\lambda_{k-1}})J_{n,z}(\lambda_{k-1})Y_n(-\lambda_{k-1})|Y_n(\lambda_{j-1})|^2\\
&+\psi(e^{i\lambda_{k-1}})J_{n,z}(-\lambda_{k-1})Y_n(\lambda_{k-1})|Y_n(\lambda_{j-1})|^2\\
&+|Y_n(\lambda_{j-1})|^2|Y_n(\lambda_{k-1})|^2.
\end{aligned} \tag{6.11}$$

For the sake of brevity we only discuss three cases in detail, as all other terms in (6.11) can be estimated by similar calculations. Following a straightforward but tedious argument [observing the representation of $Y_n(\lambda_k)$ in (6.2) and the independence of the random variables Z_t] we have

$$\begin{aligned}
E|Y_n(\lambda_k)|^8 \;\leq\; & \frac{1}{n^4}\Big(\sum_{j=-\infty}^{\infty}|\psi_j|\big(a_1|j|E|Z_1|^8 + a_2|j|^2E|Z_1|^6E|Z_1|^2\\
&+a_3|j|^3(E|Z_1|^4)^2 + a_4|j|^3E|Z_1|^4(E|Z_1|^2)^2 + a_5|j|^4(E|Z_1|^2)^4\big)^{\frac{1}{8}}\Big)^8 = O(n^{-4})
\end{aligned}$$

with appropriate constants $a_1,\dots,a_5$. This yields by (6.10) and Hölder's inequality

$$\frac{1}{n}\sum_j\sum_k E\Big(\tilde{I}_n(\lambda_j)\tilde{I}_n(\lambda_k)|Y_n(\lambda_{j-1})|^2|Y_n(\lambda_{k-1})|^2\Big) = O(n^{-1})$$

for the sum corresponding to the last term in (6.11). A similar calculation gives for the sum corresponding to the fifth term in (6.11)

$$\begin{aligned}
&\frac{1}{n}\sum_j\sum_k E\Big(\Big|\tilde{I}_n(\lambda_j)\tilde{I}_n(\lambda_k)\Psi(e^{-i\lambda_{j-1}})J_{n,z}(\lambda_{j-1})Y_n(-\lambda_{j-1})|Y_n(\lambda_{k-1})|^2\Big|\Big)\\
\leq\;& \frac{1}{n}\sum_j\sum_k\big(E|\tilde{I}_n(\lambda_j)|^4\big)^{\frac14}\big(E|\tilde{I}_n(\lambda_k)|^4\big)^{\frac14}|\Psi(e^{-i\lambda_{j-1}})|\big(E(J_{n,z}(\lambda_{j-1})Y_n(-\lambda_{j-1}))^4\big)^{\frac14}\big(E|Y_n(\lambda_{k-1})|^8\big)^{\frac14}\\
=\;& O(n^{-\frac12}),
\end{aligned}$$

where we have used the estimate

$$\Big(E\big(J_{n,z}(\lambda_{j-1})Y_n(-\lambda_{j-1})\big)^4\Big)^{\frac{1}{4}} \leq \big(E(J_{n,z}(\lambda_{j-1}))^8\big)^{\frac{1}{8}}\big(E|Y_n(-\lambda_{j-1})|^8\big)^{\frac{1}{8}} = O(n^{-\frac{1}{2}})$$

in the last step. The sums corresponding to the sixth, seventh and eighth term can be treated similarly. For a treatment of the sums corresponding to the first four terms in (6.11) we consider exemplarily the first one and use the decomposition (6.5). Since similar estimates hold for the other quantities as well, we only discuss terms involving H_n^+ as before. Thus our focus lies on

$$B_n = \frac{1}{n}\sum_j\sum_k \tilde{I}_n(\lambda_j)\tilde{I}_n(\lambda_k)\Psi(e^{-i\lambda_{j-1}})J_{n,z}(\lambda_{j-1})H_n^+(-\lambda_{j-1})\Psi(e^{-i\lambda_{k-1}})J_{n,z}(\lambda_{k-1})H_n^+(-\lambda_{k-1}).$$

Using (2.10) and (6.6) we have

$$\begin{aligned}
B_n = {} & \frac{1}{n^3}\sum_{c=-\infty}^{\infty}\sum_{d=-\infty}^{\infty}\sum_{u=1}^{n}\sum_{v=1}^{n}\sum_{m=-\infty}^{\infty}\sum_{t=1}^{n}\sum_{l=0}^{\infty}\sum_{r=1}^{l}\sum_{f=-\infty}^{\infty}\sum_{g=-\infty}^{\infty}\sum_{a=1}^{n}\sum_{b=1}^{n}\sum_{p=-\infty}^{\infty}\sum_{s=1}^{n}\sum_{q=0}^{\infty}\sum_{w=1}^{q}\psi_c\psi_d\psi_m\psi_l\psi_f\psi_g\psi_p\psi_q \\
& \times E\Big(Z_uZ_vZ_t(Z_{r-l}-Z_{n+r-l})Z_aZ_bZ_s(Z_{w-q}-Z_{n+w-q})\Big) \\
& \times\frac{1}{n^2}\sum_j\sum_k e^{-i\lambda_j(c-d+u-v+t+m-r)}e^{-i\lambda_k(f-g+a-b+s+p-w)}e^{i\lambda_1(t+m-r+s+p-w)}.
\end{aligned}$$

Because of the independence of the random variables Z_t, most of the terms in this sum will vanish. The dominating term is the sum obtained for the case where there are four different pairs of equal indices, and we will show in the following that this sum is of order $O(n^{-1})$. All other combinations will lead to terms which are of smaller order.
Let us discuss the two cases

$$u = v,\quad t = n + r - l,\quad a = b,\quad s = n + w - q$$

and

$$u = v,\quad t = b,\quad a = s,\quad r - l = w - q$$

only, as the main principle is already visible in these two situations. The corresponding term within B_n will be called $B_{n,1}$ and $B_{n,2}$, and we have

$$\begin{aligned}
E[B_{n,1}] \leq {} & C\sigma^8\frac{1}{n}\sum_{c=-\infty}^{\infty}\sum_{d=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\sum_{f=-\infty}^{\infty}\sum_{g=-\infty}^{\infty}\sum_{p=-\infty}^{\infty}\sum_{l=0}^{\infty}\sum_{r=1}^{l}\sum_{q=0}^{\infty}\sum_{w=1}^{q}\psi_c\psi_d\psi_m\psi_l\psi_f\psi_g\psi_p\psi_q \\
& \sum_{t=1}^{n}\sum_{s=1}^{n}\frac{1}{n^2}\Big(\sum_j e^{-i\lambda_j(c-d+t+m-r)}\Big)\Big(\sum_k e^{-i\lambda_k(f-g+s+p-w)}\Big)e^{i\lambda_1(t+m-r+s+p-w)}
\end{aligned}$$

as well as

$$\begin{aligned}
E[B_{n,2}] \leq {} & C\sigma^8\frac{1}{n^2}\sum_{c=-\infty}^{\infty}\sum_{d=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\sum_{f=-\infty}^{\infty}\sum_{g=-\infty}^{\infty}\sum_{p=-\infty}^{\infty}\sum_{l=0}^{\infty}\sum_{r=1}^{l}\sum_{q=0}^{\infty}\sum_{w=1}^{q}\psi_c\psi_d\psi_m\psi_l\psi_f\psi_g\psi_p\psi_q \\
& \sum_{t=1}^{n}\sum_{s=1}^{n}\frac{1}{n^2}\Big(\sum_j e^{-i\lambda_j(c-d+t+m-r)}\Big)\Big(\sum_k e^{-i\lambda_k(f-g+2s-t+p-w)}\Big)e^{i\lambda_1(t+m-r+s+p-w)}.
\end{aligned}$$

Since both $1 \leq t \leq n$ and (6.7) hold, we conclude as in the previous proof that there is actually only one index t that gives a non-zero term in $E[B_{n,1}]$. Similar derivations for s and $B_{n,2}$ plus the same argument as in (6.8) yield

$$|E[B_{n,1}]| \leq C\frac{\sigma^8}{n} \sum_{c=-\infty}^{\infty} \sum_{d=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \sum_{f=-\infty}^{\infty} \sum_{g=-\infty}^{\infty} \sum_{p=-\infty}^{\infty} \sum_{l=0}^{\infty} \sum_{r=1}^{l} \sum_{q=0}^{\infty} \sum_{w=1}^{q} \Big|\psi_c\psi_d\psi_m\psi_l\psi_f\psi_g\psi_p\psi_q\Big| = O(n^{-1})$$

as well as $|E[B_{n,2}]| = O(n^{-2})$. Putting things together, we obtain $|E[B_n]| = O(n^{-1})$, and a similar calculation for the sum corresponding to the remaining terms in (6.11) yields (6.9) and therefore assertion (2.9).

In the multivariate case, related arguments yield

$$\Big|\frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} I_{n,ij}(\lambda_k) I_{n,ji}(\lambda_{k-1}) - \frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} \tilde{I}_{n,ij}(\lambda_k) \tilde{I}_{n,ji}(\lambda_{k-1})\Big| = o_P(1)$$

as an analogue of (2.9) plus the claim in (3.10) for arbitrary i and j. Note in particular that the symmetry argument from (6.8) works in this context as well. □

Acknowledgements. The authors would like to thank Martina Stein, who typed parts of this manuscript with considerable technical expertise. This work has been supported in part by the Collaborative Research Center "Statistical modeling of nonlinear dynamic processes" (SFB 823) of the German Research Foundation (DFG).

References

Anderson, T. W. (1993). Goodness of fit tests for spectral distributions. *Annals of Statistics*, 21:803–847.

Bartlett, M. S. (1955). *An Introduction to Stochastic Processes with Special Reference to Methods and Applications*. Cambridge University Press, London.

Box, G. E. P. and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive integrated moving average time series models. *Journal of the American Statistical Association*, 65:1509–1526.

Brockwell, P. and Davis, R. (1991). *Time Series: Theory and Methods*. Springer Verlag, New York.

Carmona, R. A. and Wang, A. (1996). Comparison tests for the spectra of dependent multivariate time series. In Adler, R. J., Müller, P., and Rozovskii, B. L., editors, *Stochastic Modeling in Physical Oceanography*, pages 69–88. Birkhäuser Inc., Boston.

Chen, H. and Romano, J. (1999). Bootstrap goodness of fit tests in the frequency domain. *Journal of Time Series Analysis*, 20:619–654.

Coates, D. S. and Diggle, P. J. (1986). Test for computing two estimated spectral densities. *Journal of Time Series Analysis*, 7:7–20.

Dahlhaus, R. (1985). On the asymptotic distribution of bartlett's U_p-statistic. *Journal of Time Series Analysis*, 6:213–227.

Dette, H. and Hildebrandt, T. (2010). A note on testing hypotheses for stationary processes in the frequency domain. *Submitted for publication, http://www.ruhr-uni-bochum.de/mathematik3/en/research/index.html*.

Dette, H. and Paparoditis, E. (2009). Bootstrapping frequency domain tests in multivariate time series with an application to comparing spectral densities. *Journal of the Royal Statistical Society: Series B*, 71:831–857.

Dette, H. and Spreckelsen, I. (2000). A test for randomness against ARMA alternatives. *Stochastic Processes and their Applications*, 89:131–139.

Dette, H. and Spreckelsen, I. (2003). A note on a specification test for time series models based on spectral density estimation. *Scandinavian Journal of Statistics*, 30:481–491.

Diggle, P. J. and Fisher, N. I. (1991). Nonparametric comparison of cumulative periodograms. *Applied Statistics*, 40:423–434.

Eichler, M. (2008). Testing nonparametric and semiparametric hypotheses in vector stationary processes. *Journal of Multivariate Analysis*, 99:968–1009.

Hannan, E. J. (1970). *Multiple Time Series*. Wiley, New York.

Kakizawa, Y., Shumway, R. H., and Taniguchi, M. (1998). Discrimination and clustering for multivariate time series. *Journal of the American Statistical Association*, 93:328–340.

Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65:297–303.

Matsuda, Y. and Yajima, Y. (2004). On testing for separable correlations of multivariate time series. *Journal of Time Series Analysis*, 25:501–528.

Mokkadem, A. (1997). A measure of information and its applications to test for randomness against ARMA alternatives and to goodness-of-fit test. *Stochastic Processes and their Applications*, 72:145–159.

Monti, A. C. (1994). A proposal for a residual autocorrelation test in linear models. *Biometrika*, 81:776–780.

Orey, S. (1958). A central limit theorem for m-dependent random variables. *Duke Mathematical Journal*, 52:543–546.

Paparoditis, E. (2000). Spectral density based goodness-of-fit tests for time series models. *Scandinavian Journal of Statistics*, 27:143–176.

Peña, D. and Rodriguez, J. (2002). A powerful portmanteau test of lack of fit for time series. *Journal of the American Statistical Association*, 97:601–610.

Swanepoel, J. W. H. and van Wyk, J. W. J. (1986). The comparison of two spectral density functions using the bootstrap. *Journal of Statistical Computation and Simulation*, 24:271–282.

Taniguchi, M. and Kondo, M. (1993). Non-parametric approach in time series analysis. *Journal of Time Series Analysis*, 14:397–408.

Taniguchi, M., Puri, M. L., and Kondo, M. (1996). Nonparametric approach for non-Gaussian vector stationary processes. *Journal of Multivariate Analysis*, 56:259–283.

Zhang, G. and Taniguchi, M. (1994). Discriminant analysis for stationary vector time series. *Journal of Time Series Analysis*, 15:117–126.